

(12) **Österreichische Patentanmeldung**

(21) Anmeldenummer: A 708/2010  
(22) Anmeldetag: 28.04.2010  
(43) Veröffentlicht am: 15.11.2011

(51) Int. Cl. : **G06F 17/17** (2006.01)  
**C12Q 1/68** (2006.01)

(73) Patentanmelder:  
AIT AUSTRIAN INSTITUTE OF TECHNOLOGY GMBH  
A-1220 WIEN (AT)

(54) **VERFAHREN ZUR APPROXIMATION EINER SCHMELZKURVE**

(57) die Erfindung betrifft ein Verfahren zur Approximation einer Schmelzkurve von in einer Probe befindlichen Moleküle, insbesondere Biomolekülen, wie z.B. DNA, RNA-, PNA-, LNA-Molekülen oder Proteinen, die an Fängermolekülen anhaften, basierend auf einer Vielzahl von Messdatensätzen umfassend jeweils einen Messwert  $I$ , der proportional zur Konzentration der an einem Fängermolekül anhaftenden zu messenden Proteine ist, sowie einen zugehörigen gemessenen Temperaturwert  $T$ . Erfindungsgemäß ist vorgesehen, dass die vorgegebenen Messwerte mittels eines Fitting-Verfahrens an eine den Verlauf des Messwerts  $I$  für eine vorgegebene Temperatur  $T$  bezeichnende Kurve der

$$\text{Form} \quad I(T) = \frac{k_a (k_{Cy} - T)}{e^{\frac{k_a (T - k_{Tm})}{k_b}} + 1} - k_{HT} * T + k_{const}$$

angepasst werden, und wobei die Fitting-Parameter  $k_a$ , und  $k_b$ ,  $k_{Cy}$ ,  $k_{HT}$ ,  $k_{Tm}$  und  $k_{const}$  im Zuge des Fitting-Verfahrens an die aufgenommenen Messdatensätze angepasst werden.

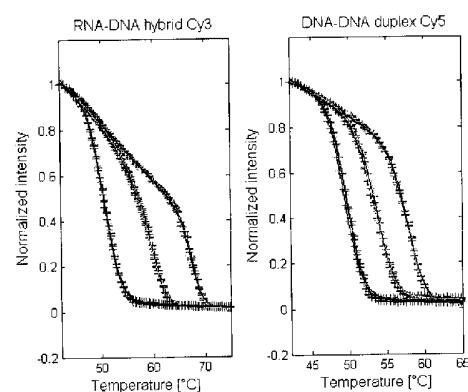


Fig. 1

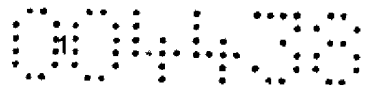
# Zusammenfassung:

Die Erfindung betrifft ein Verfahren zur Approximation einer Schmelzkurve von in einer Probe befindlichen Moleküle, insbesondere Biomolekülen, wie z.B. DNA, RNA-, PNA-, LNA-Molekülen oder Proteinen, die an Fängermolekülen anhaften, basierend auf einer Vielzahl von Messdatensätzen umfassend jeweils einen Messwert I, der proportional zur Konzentration der an einem Fängermolekül anhaftenden zu messenden Proteine ist, sowie einen zugehörigen gemessenen Temperaturwert T. Erfindungsgemäß ist vorgesehen, dass die vorgegebenen Messwerte mittels eines Fitting-Verfahrens an eine den Verlauf des Messwerts I für eine vorgegebene Temperatur T bezeichnende Kurve der

Form 
$$I(T) = \frac{k_a (k_{Cy} - T)}{e^{k_b (T - k_{Tm})} + 1} - k_{HT} * T + k_{const}$$
 angepasst werden, und wobei die Fitting-

Parameter  $k_a$ , und  $k_b$ ,  $k_{Cy}$ ,  $k_{HT}$ ,  $k_{Tm}$  und  $k_{const}$  im Zuge des Fitting-Verfahrens an die aufgenommenen Messdatensätze angepasst werden.

(Fig. 1)



Die Erfindung betrifft ein Verfahren zur Approximation einer Schmelzkurve gemäß dem Oberbegriff des Patentanspruchs 1.

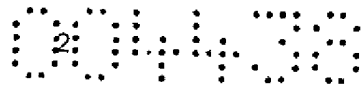
Die Erfindung wird im Bereich der automatisierten Analyse zur Schmelzkurvenbestimmung DNA, RNA, PNA, LNA, Proteinen usw. gewerblich eingesetzt.

Hintergrund der Erfindung ist die automatisierte Aufzeichnung von Schmelzkurven. In der momentan verfügbaren Microarray-Technologie zur Beobachtung und Messung von Proben mit DNA, RNA, PNA, LNA Proteinen, usw. können Bilder von Microarrays während des Ausschmelzens ohne weiteres aufgenommen werden. Problematisch stellt sich hingegen die Analyse der so aufgenommenen Messwerte bzw. Bilder dar, da eine große Datenmenge für die einzelnen Messpunkte eines Microarrays vorliegt und eine Analyse der aufgenommenen Daten in relativ kurzer Zeit erfolgen muss.

Im Stand der Technik sind einige Verfahren zur allgemeinen Beschreibung temperaturabhängigen Reaktionen bekannt, die jedoch nicht speziell an die Verwendung bei Proben mit DNA, RNA, PNA, LNA oder Proteinen angepasst ist. Hierbei ist insbesondere problematisch, dass die Einflüsse von Oberflächeneffekten am Probenträger und sterischen Effekten der untersuchten Moleküle, Restriktionen in Bezug auf die Entropie sowie elektrostatische Effekte nicht berücksichtigt werden. Zudem besteht die Herausforderung eines sehr geringen Signal-Hintergrund-Abstand sowie des Rauschen durch die Segmentierung sowie weitere Messartefakte.

Bei den bisherigen bekannten Verfahren werden z.B. zwei Grundlinien herangezogen, die an die ermittelte Schmelzkurve angepasst werden, wobei die Anpassung außerhalb der eigentlichen Messung erfolgt. Wesentlicher Nachteil dieses Verfahrens ist, dass die Anpassung nicht mittels eines automatisierten Verfahrens sondern ausschließlich mit menschlichem Zutun erfolgt. Zudem ist es mit derartigen Verfahren nicht möglich, mehrere Proben mit sehr unterschiedlicher Schmelztemperatur innerhalb eines Messschritts zu behandeln. Da bei der Vorgabe der Grundlinien keine Parameter vorgegeben werden können, ist auch eine Qualitätsüberprüfung der vorgegebenen Grundlinien nur durch visuelle Kontrolle möglich.

Ferner sind ableitungsbasierte Verfahren mit unterschiedlichen Filter-Algorithmen und modellbasierten Fitting-Algorithmen bekannt, die bei lösungsbasierten Proben verwendet werden. So kann beispielsweise ein digitaler Tiefpassfilter zur Filterung der aufgenommenen Messwerte herangezogen werden. Ein derartiges Verfahren eignet sich jedoch nur bedingt zur automatisierten Analyse der aufgenommenen Messdatensätze, da eine händische Einstellung des Tiefpassfilters erforderlich ist.



Ferner ist das Savitzky-Golay Verfahren bekannt, das bei sehr verrauschten Funktionen die Form der Funktion gut erhält. Obwohl dieses Verfahren prinzipiell sehr gute Ergebnisse liefert, bietet es kaum Vorteile gegenüber dem Tiefpassfilter, da für die Form der Schmelzkurven kein Modell für die zu erwartende Kurvenform vorliegt.

Aufgabe der Erfindung ist es somit, ein Verfahren zur Verfügung zu stellen, das eine automatisierte Bestimmung von Schmelzkurven erlaubt und einen menschlichen Eingriff in die Auswertung überflüssig macht.

Die Erfindung löst die Aufgabe bei einem Verfahren zur Approximation einer Schmelzkurve von in einer Probe befindlichen DNA-, RNA-, PNA-, LNA-Molekülen oder Proteinen, die an Fängermolekülen anhaften, basierend auf einer Vielzahl von Messdatensätzen umfassend jeweils einen Messwert  $I$ , der proportional zur Konzentration der an einem Fängermolekül anhaftenden zu messenden Nucleinsäuren oder Proteine ist, sowie einen zugehörigen gemessenen Temperaturwert  $T$ . Erfindungsgemäß ist vorgesehen, dass die vorgegebenen Messwerte mittels eines Fitting-Verfahrens an eine den Verlauf des Messwerts  $I$  für eine vorgegebene Temperatur  $T$  bezeichnende Kurve der Form

$$I(T) = \frac{k_a(k_{cy} - T)}{e^{k_b(T - k_{tm})} + 1} - k_{HT} * T + k_{const} \text{ angepasst werden, und wobei die Fitting-}$$

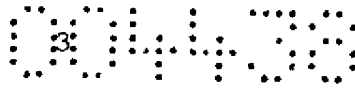
Parameter  $k_a$ , und  $k_b$ ,  $k_{cy}$ ,  $k_{HT}$ ,  $k_{tm}$  und  $k_{const}$  im Zuge des Fitting-Verfahrens an die aufgenommenen Messdatensätze angepasst werden.

Mit diesem Verfahren können Schmelzkurven besonders einfach und ressourcensparend angenähert werden. Durch die nach wenigen Messungen mögliche Prognose kann die Detektion auf Vorliegen von bestimmten Substanzen bzw. Molekülen drastisch verkürzt werden.

Eine bevorzugte Ausführungsform der Erfindung sieht vor, dass sämtliche Messwerte einer linearen Transformation unterzogen werden, die diese auf Werte innerhalb des Intervalls  $[0, 1]$  abbildet, wobei insbesondere der Wert  $I(T_{min})$  der Anfangstemperatur auf 1 und der Wert  $I(T_{max})$  auf null abgebildet wird.

Hierdurch wird die Vergleichbarkeit unterschiedlicher Messungen mit unterschiedlichen Konzentrationen von zu prüfenden Molekülen verbessert. Zudem kann das Verfahren somit unabhängig von der Konzentration der Markermoleküle betrieben werden.

Weiters kann vorgesehen sein, dass die Temperatur der Probe kontinuierlich erhöht wird, und die Probe gegebenenfalls bestrahlt und/oder mit Markermolekülen versetzt wird, und die Messdatensätze aufgezeichnet werden, indem in vorgegebenen Zeitabständen



Messungen der Messwerte gemeinsam mit der zum selben Zeitpunkt gemessenen Temperatur (T) vorgenommen werden.

Hierdurch wird die Genauigkeit der Messwerte erhöht und die Prognosequalität verbessert.

Ferner ist es vorteilhaft, dass laufend jeweils nach der Bestimmung einer Anzahl von Messdatensätzen jeweils ein Fitting-Schritt durchgeführt wird und basierend auf den bisher ermittelten Ergebnissen jeweils eine verbesserte Fitting-Kurve erstellt wird.

Dieses Vorgehen verbessert die Prognosequalität und erleichtert die Abschätzung der Messergebnisse für den Benutzer.

Vorteilhafterweise kann vorgesehen sein, dass die Probe in einen Probenträger verbracht wird, wobei an einer der Begrenzungsflächen des Probenträgers Fängermoleküle angeordnet sind, die eine vorgegebene Art von Molekülen binden, wobei die Anzahl der an den Fängermolekülen befindlichen Moleküle, z.B. von DNA, RNA, PNA, LNA oder Proteine als Messwert I bestimmt werden.

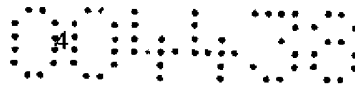
Dieses Vorgehen verbessert die Handhabung bei der Durchführung der Messung.

Eine bevorzugte Weiterbildung der Erfindung sieht vor, dass als Messwert der von der Probe reflektierte und/oder, insbesondere durch Lumineszenz, vorzugsweise Fluoreszenz, abgestrahlte Lichtintensität herangezogen wird und/oder als Messwert die elektrische Leitfähigkeit, die Magnetisierbarkeit, die Impedanz bei einer vorgegebenen Frequenz, insbesondere zwischen 1 Hz und  $10^7$  Hz, oder die Änderung des Brechungsindex herangezogen wird und/oder als Messwert I die Gewichtszunahme bedingt durch das Anhaften der Moleküle an den Fängermolekülen herangezogen wird.

Hierdurch werden einfache und genaue Möglichkeiten zur Bestimmung der dem Verfahren zugrundeliegenden Messwerte dargestellt.

Zudem kann vorgesehen sein, dass der Probenträger an zumindest einer Seite von einem Microarray begrenzt ist, und die Bestimmung des Messwerts durch Aufnahme eines Digitalbildes des Microarrays durchgeführt wird.

Mit einem derartigen Vorgehen kann eine Vielzahl unterschiedlicher Messungen, z. B. eine Messung unterschiedlicher Proben mit denselben Fängermolekülen oder eine Messung derselben Probe mit unterschiedlichen Fängermolekülen einfach und effizient durchgeführt werden.



Weiters ist es vorteilhaft, dass jedem Rasterpunkt des Microarrays vorab ein oder mehrere Pixel des Digitalbilds zugeordnet werden und aus den Lichtintensitätswerten der jeweiligen Pixel die jeweiligen Messwerte  $I$  der in den Rasterpunkten des Microarrays befindlichen Proben bestimmt werden und dass für jeden Rasterpunkt des Microarrays laufend separate Messungen und Fitting-Verfahrensschritte durchgeführt werden.

Hierdurch wird zusätzlich die Aufnahme einer Vielzahl von Messwerten erleichtert und effizient durchgeführt.

Vorteilhafterweise kann vorgesehen sein, dass die Temperatur  $T$  mit einer Rate von 0,1 bis 5°C/min, insbesondere 0,8 bis 1,5°C/min erhöht wird.

Mit diesem Vorgehen kann die Schmelzkurve mit wenigen Messwerten mit ausreichender Genauigkeit bestimmt werden.

Eine bevorzugte Ausführungsform der Erfindung sieht vor, dass zur Bestimmung der Fitting-Parameter  $k_a$ , und  $k_b$ ,  $k_{Cy}$ ,  $k_{HT}$ ,  $k_{Tm}$  und  $k_{const}$  ein Gauß-Newton-Verfahren verwendet wird.

Mit diesem Vorgehen wird eine der Schmelzkurve besonders gut angepasste Kurvenform und hervorragende Prognosen erzielt.

Die Erfindung wird anhand einiger Ausführungsbeispiele ohne Einschränkung der Allgemeinheit des erfinderischen Gedankens dargestellt.

Fig. 1 zeigt Beispiele für normalisierte Schmelzkurven.

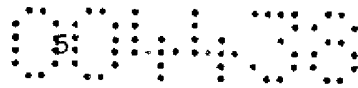
Fig. 2 zeigt nicht normalisierte Schmelzkurven und deren Modellierung.

Fig. 3 zeigt das erfindungsgemäße Verfahren mit einer eingeschränkten Anzahl von Messpunkten.

Fig. 4 zeigt die Prognose einer Schmelzkurve.

Fig. 5 zeigt die Extrapolation von Kurven, wobei lediglich Messwerte unterhalb des Schmelzpunktes vorliegen.

Die Messwerte, die im Zuge des erfindungsgemäßen Verfahrens weiterverarbeitet werden, werden typischerweise während einer Reaktion aufgezeichnet, bei der sich die Probe in einem Probenträger befindet. Die Probe enthält insbesondere Proteine, DNA, RNA, LNA oder PNA. An einer Begrenzungsfläche des Probenträgers ist eine Anzahl von Fängermolekülen angeordnet, die fest an der Begrenzungsfläche anhaften. Diese Fängermoleküle sind Moleküle, die komplementär zu den zu untersuchenden Molekülen aufgebaut sind, sodass die zu untersuchenden Moleküle leicht an den Fängermolekülen anhaften. Bei den zu untersuchenden Molekülen handelt es sich vorteilhafterweise um



Biomoleküle wie DNA, RNA, PNA, LNA oder Proteine. Je besser die Übereinstimmung der beiden Moleküle, desto besser haften das zu untersuchende Molekül und das Fängermolekül aneinander und desto höhere Temperaturen sind erforderlich, um die Moleküle voneinander zu trennen. Der Vorgang der Trennung wird als Schmelzen bezeichnet. Eine Schmelzkurve ist demgemäß eine Kurve, die für eine Anzahl von Temperaturen die jeweils an den Fängermolekülen anhaftenden Moleküle einer Probe angibt.

Um z. B. das Vorhandensein von bestimmten Viren in einer Probe zu analysieren, werden Fängermoleküle verwendet, die eine dem Virus komplementäre Struktur aufweisen, sodass typische Nukleinsäuresequenzen der Viren an den Fängermolekülen anhaften. Sind in der Probe beispielsweise auch andere Viren enthalten, so haften diese Nukleinsäuresequenzen ebenfalls an den Fängermolekülen an, jedoch mit verringerter Bindungsenergie, sodass sich diese bei hohen Temperaturen von den gesuchten Sequenzen wieder von den Fängermolekülen lösen. Da die Bindungsenergie je nach Art des zu untersuchenden Moleküls unterschiedlich ist, sind Kalibriermessungen mit positiven Proben vorteilhaft, um Aussagen über das Vorhandensein z. B. eines Virus in einer bestimmten Probe treffen zu können.

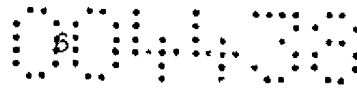
Das vorliegende Verfahren ist jedoch nicht auf die Detektion von Viren beschränkt sondern kann für sämtliche Moleküle verwendet werden, für die komplementäre Fängermoleküle zur Verfügung stehen. Dies ist zumindest für DNA, RNA, PNA, LNA und Proteine der Fall.

Um die Anzahl der an den Fängermolekülen anhaftenden zu untersuchenden Moleküle feststellen zu können, besteht die Möglichkeit, die zu untersuchenden Moleküle zu markieren, sodass diese bei Bestrahlung mit Licht bestimmter Wellenlänge nachleuchten (phosphoreszieren) bzw. bei einer anderen Wellenlänge Licht aussenden (fluoreszieren).

Bei der Bestimmung der Schmelzkurve einer Probe wird die Probe auf einen Probenträger aufgebracht. Als Probe werden beispielsweise DNA- oder RNA- Moleküle, oder Proteine herangezogen und auf den Probenträger aufgebracht, die einer vorgegebenen Reaktion unterliegen. Diese Reaktion kann je nach Art zwischen wenigen Minuten und wenigen Tagen dauern.

Zur gleichzeitigen Analyse von mehreren Proben mit derselben Analyseflüssigkeit können Microarrays verwendet werden. Nach dem Aufbringen der Probe auf dem Probenträger wird eine Analyseflüssigkeit mit Markermolekülen auf die Proben aufgebracht.

Wird die Probe mit Licht, insbesondere Licht im Wellenlängenbereich zwischen 500nm und 700 nm bestrahlt, fluoreszieren die Markermoleküle und die von den Markermolekülen reflektierte Lichtintensität ist annähernd proportional zur Anzahl der markierten Moleküle. Die Intensität wird mittels eines Photodetektors ermittelt und liegt in



quantisierter Form vor. Dabei ist es nicht erforderlich, den Wert zu kalibrieren, ein dimensionsloser quantisierter Messwert  $I$  ist für das im folgenden angeführte Verfahren ausreichend.

Anschließend wird die Probe kontinuierlich mit vorgegebenem Temperaturanstieg mit einer Rate von etwa 0,1 bis 5°C/min, insbesondere 0,8 bis 1,5°C/min erhitzt. Die in den Proben vorliegenden zu untersuchenden Moleküle haften zunächst an den Fängermolekülen an. Durch die Erwärmung werden die Moleküle mit geringerer Bindungsenergie aus den Fängermolekülen gelöst, wobei die Moleküle mit geringerer Bindungsenergie bei niedrigeren Temperaturen gelöst werden.

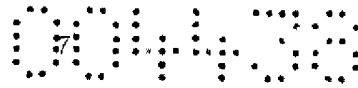
Auf Grund der Erhitzung werden die Moleküle angeregt und die Bindungen zwischen den einzelnen langkettigen Molekülen werden aufgebrochen. Während bei geringeren Temperaturen nur eine kleine Anzahl von Bindungen gelöst wird, werden bei größeren Temperaturen durch die höheren thermischen Bewegungen der Moleküle mehr Bindungen gelöst. Die Anzahl der an die Probe gebundenen Markermoleküle verringert sich und bewirkt eine Reduktion der Fluoreszenz der Probe bis hin zu einem Minimum.

Im dargestellten Ausführungsbeispiel werden die Lichtintensitäten der einzelnen Proben in Abständen von einigen Sekunden bis zu einer Minute aufgenommen und abgespeichert. Zusätzlich wird für jeden der Aufnahmezeitpunkte jeweils die Temperatur der Probe bzw. der Proben bestimmt. Dabei kann für eine Mehrzahl von Proben stets dieselbe Temperatur angenommen werden. Nach der Durchführung einer Messung erhält man jeweils einen Messdatensatz mit einer ermittelten Intensität der Fluoreszenz sowie eine zugehörige Temperatur. Die Zeitpunkte der Messung sind für die folgenden Betrachtungen ohne Belang.

Alternativ können anstelle der Lichtintensität der Fluoreszenz auch elektrische, mechanische, chemische oder magnetische Eigenschaften der Biomoleküle gemessen werden, die sich bei einer Bindung mit den zu messenden Molekülen ändern. So ändern sich beispielsweise die elektrische Leitfähigkeit sowie die Magnetisierbarkeit der Moleküle abhängig von der Menge der an den Fängermolekülen anhaftenden Molekülen. Weiters kann alternativ die Impedanz bei einer vorgegebenen Frequenz, insbesondere zwischen 1 Hz und  $10^7$  Hz. herangezogen werden, um die Anzahl der an den Fängermolekülen anhaftenden Moleküle zu bestimmen.

Eine weitere Möglichkeit, die Anzahl der an den Fängermolekülen anhaftenden Moleküle zu bestimmen, liegt darin, das Gewicht bzw. die Masse der Gesamtheit der Fängermoleküle sowie der daran anhaftenden Moleküle zu bestimmen. Dies kann beispielsweise durch Piezo-Kristalle erfolgen, wobei deren Resonanzfrequenz sich abhängig von der Masse der Fängermoleküle ändert. Alternativ dazu kann auch das Gewicht der Fängermoleküle und der an ihnen anhaftenden Moleküle bestimmt werden





und die am Piezo-Kristall anliegenden Gleichspannung ermittelt werden, die proportional der Masse der Fängermoleküle sowie der daran anhaftenden Moleküle ist.

Im Zuge des Fitting-Verfahrens wird eine Ausgleichskurve gesucht, die diese Messdatensätze mit besonders geringer Fehlerrate approximiert. Es wird dabei eine

grundsätzliche Kurvenform der Formel  $I(T) = \frac{k_a(k_{cy} - T)}{e^{k_b(T - k_{tm})} + 1} - k_{HT} * T + k_{const}$  gewählt und

die Fitting-Parameter  $k_a$ , und  $k_b$ ,  $k_{cy}$ ,  $k_{HT}$  und  $k_{const}$  im Zuge des Fitting-Verfahrens an die aufgenommenen Messdatensätze angepasst.

Es ist wünschenswert, den Verlauf der Schmelzkurve so darzustellen, dass der Temperaturverlauf der Konzentration eines der Bestandteile der Probe vorliegt. Dabei ist die eigentliche Konzentration des Bestandteils der Probe weniger von Belang als die Kurvenform des Temperaturverlaufs, der im Folgenden als Schmelzkurve bezeichnet wird. Als Normierung der Kurvenform wird eine lineare Transformation der Konzentrationen in das Intervall  $[0, 1]$  bezeichnet, wobei das Maximum der Konzentrationen auf 1 und das Minimum der Konzentrationen auf 0 abgebildet wird. Die Schmelzkurve wird als Kurve bezüglich eines Koordinatensystems dargestellt, wobei dessen Ordinate ein normalisierter Messwert  $I$  der Dimension 1 darstellt. Auf der Abszisse des Koordinatensystems ist die Temperatur aufgetragen. Die Schmelzkurve  $I(T)$  ist, wie in Fig. 1 und 2 dargestellt, in das Koordinatensystem eingetragen. Auf Grund der Normierung der Schmelzkurve, die für das Optimierungsverfahren als zusätzliche Option vorgegeben wird, beginnt die Kurve  $I(T)$  bei einer vorgegebenen Anfangstemperatur  $T_{min}$  beim Wert 1 und endet bei einer vorgegebenen Endtemperatur  $T_{max}$  beim Wert 0.

Die Parameter  $k_a$  und  $k_b$  sind Normalisierungsparameter, deren Auswahl im Zuge des Fitting-Verfahrens dahingehend festgelegt wird, dass der Anfangswert der Schmelzkurve  $I(T_{min})$  bei 1 und der Endwert der Schmelzkurve  $I(T_{max})$  bei 0 liegt.

Der zusätzliche Term  $-k_{HT} * T + k_{const}$  geht auf den Umstand ein, dass selbst wenn keinerlei Bindungen mehr zwischen den Markermolekülen und den Proben vorliegen, eine temperaturabhängige Fluoreszenz messbar ist. Dabei wird eine negative lineare Temperaturabhängigkeit der Fluoreszenz angenommen. Dabei handelt es sich um einen Artefakt der durch die unspezifische Bindung der Markermoleküle an beliebige Moleküle.

Der Fitting-Parameter  $k_{cy}$  dient zur Anpassung der Schmelzkurve an die lineare Abhängigkeit der Färbung der Markermoleküle von der Temperatur.

Der nichtlineare Teil der grundlegenden Kurvenform wird durch eine Exponentialfunktion mit zwei Parametern  $k_{tm}$  und  $k_b$  beschrieben. Diese Funktion ist ähnlich der Fermi-Dirac-Verteilung, die die Verteilung von einzelnen Teilchen in einem Mehrteilchensystem von identischen Fermionen wiedergibt. Obwohl es hier keine offensichtlichen Beziehungen der

physikalischen Prozesse gibt, haben Experimente eine gute numerische Übereinstimmung für die Beschreibung der Schmelzkurven gezeigt.

Die Kurve  $I(T)$  kann mittels eines Fitting-Verfahrens gemäß dem Stand der Technik, beispielsweise einem Gauß-Newton-Verfahren (Björck, A. (1996). Numerical methods for least squares problems. SIAM, Philadelphia. ISBN 0-89871-360-9), ermittelt werden. Wie aus der Fig. 1 ersichtlich, sind die Abweichungen von den einzelnen Messwerten sehr gering.

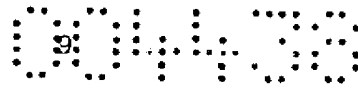
Fig. 1 zeigt ein Beispiel für normalisierte Schmelzkurven für RNA-DNA und DNA-DNA Duplexe. Die Kreuze zeigen Messwerte, die Linien die nach dem Verfahren ermittelten Kurven an. Es ist für alle Fälle sehr gute Übereinstimmung sichtbar.

Fig. 2 zeigt ein Beispiel von gemessenen nicht normalisierten Schmelzkurven und deren Modellierung. Es werden alle experimentell bestimmten Messwerte dargestellt. Die geringe Abweichung der mit dem Verfahren ermittelten Funktion mit den experimentellen Werten ist unmittelbar ersichtlich. Dargestellt ist der Intensitätsverlauf über die Temperatur einer Schmelzkurvenbestimmung von vier unterschiedlichen Arten von RNA-Molekülen auf demselben Fängermolekül.

Fig. 3 zeigt die Anwendung des Verfahrens auf eine stark eingeschränkte Menge von Messwerten. Es wurde nur jeder zehnte Messpunkt für die Modellierung verwendet. Die Messwerte, die zur Bildung der Kurve  $I(T)$  verwendet worden sind, sind mit einem Stern gekennzeichnet. Trotz des verwendeten Intervalls von rund  $4^{\circ}\text{C}$  zwischen den Messwerten kann eine sehr gute Genauigkeit von weit weniger als  $1^{\circ}\text{C}$  erreicht werden. Fig. 3 liegen dieselben Reaktionen zugrunde wie der Fig. 2.

Fig. 4 zeigt Prognose einer Schmelzkurvenmessung zu höheren Temperaturen. Die durch einen Stern gekennzeichneten Messwerte wurden für den erfindungsgemäßen Algorithmus herangezogen. Die Punkte stellen die Messwerte dar, die jedoch nicht für den Algorithmus herangezogen wurden, die Kurven resultieren aus der Modellierung. Die Auswahl der Messwerte erfolgte so das die gewählten Messtemperaturen kleiner waren als die Schmelztemperatur. Dargestellt sind der Intensitätsverlauf von vier unterschiedlichen DNA-Molekülen auf demselben Fängermolekül.

Fig. 5 zeigt die Anwendung des Algorithmus für einen Datensatz für eine reduzierte Anzahl an Messwerten. Es wurden nur die mit einem Stern gekennzeichneten Messwerte



dem Algorithmus zugrundegelegt. Es wurde für alle Versuche eine Restintensität, d.h. ein zu kompensierender Offset verursacht durch die unspezifische Bindung von Fluorophoren im Zuge der Messung von 100 angenommen. Es zeigt sich, dass auch durch Messungen unterhalb der eigentlichen Schmelztemperatur, hier etwa 55°C, die thermodynamischen Parameter präzise zu ermitteln sind. Fig. 5 liegen dieselben Reaktionen zugrunde wie Fig. 4.

Im folgenden Beispiel wird die konkrete Messung einer typischen Versuchsmessung im Detail beschrieben:

Allgemein gilt für die Intensität

$$I = I_{sp} + I_{Bg} \quad (1)$$

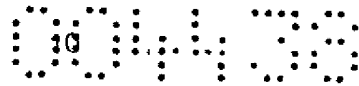
$I_{sp}$  ist die Intensität verursacht durch die spezifische Hybridisierung komplementärer Nukleinsäuren,  $I_{Bg}$  wird vom Hintergrund und unspezifischen Effekten verursacht. Bei Hybridisierung oberflächengebundener Nukleinsäuren beeinflussen viele Faktoren die Intensität  $I_{Bg}$  z.B. unspezifische Hybridisierung, Oberflächeneffekte, optische Artefakte und Prozessstreuungen (Peterson et al., 2002). Die Entwicklung im Design spezifischer Fängermoleküle ist immer noch nicht abgeschlossen (Gresham et al., 2010; Leparc et al., 2009). Weitere Verbesserungen erhofft man sich von Algorithmen zur genaueren Bestimmung des unspezifischen Hintergrundes (Furusawa et al., 2009; Hooyberghs et al., 2009; Kroll et al., 2009; Ono et al., 2008).

Beim Design von Fängermolekülen wird die Enthalpie  $\Delta H^0$  und die Entropie  $\Delta S^0$  mit der Schmelztemperatur über

$$T_m = \frac{\Delta H^0}{R \ln\left(\frac{C_t}{4}\right) + \Delta S^0} \quad (2)$$

verknüpft.  $R$  ist die Gaskonstante,  $C_t$  die Konzentration in mol. Die Messdaten stammen aus Lösungs-basierten Schmelzkurven. Mit diesen Methoden werden die komplexen Verhältnisse bei oberflächengebundenen Fängermolekülen nur unvollständig wiedergegeben (Li et al., 2008; Pozhitkov et al., 2007).

Echtzeit Datenaufnahme während Hybridisierung und Denaturierung (Bishop et al., 2007; Chagovetz and Blair, 2009; Khomyakova et al., 2008; Marcy et al., 2008) überwinden diese Probleme da die Änderung des spezifischen Signals  $I_{sp}$  in Gleichung (1) vom konstanten unspezifischen Signal  $I_{Bg}$  unterscheidbar ist. Es ist bekannt das die Hybridisierungskinetik Informationen zur Unterscheidung von spezifischer und unspezifischer Hybridisierung und zur quantitativen Analyse liefert (Bishop et al., 2008; Dai et al., 2002; Khomyakova et al., 2008). Durch die Vielfachmessung können auch



fortschrittliche Methoden zur Signalverarbeitung genutzt werden, besonders das Verwenden von physikalischen Modellierungsfunktionen ermöglicht eine klare Unterscheidung von Signal und Hintergrund. Ein guter Übersichtsartikel über Echtzeitmikroarrays ist (Chagovetz and Blair, 2009).

Schmelzkurvenanalyse wird bereits für lösungsgebundenes „high resolution melting“ (HRM) (Dahl and Guldborg, 2007; Herrmann et al., 2006; Wojdacz and Dobrovic, 2007) und für komplizierte oberflächengebundene Verfahren mit geringen Spotdichten verwendet: SPR (Xu et al., 2008; Yao et al., 2005), DASH (Howell et al., 1999; Stromqvist Meuzelaar et al., 2007). Da bei jeder Aufnahme ein Bild gespeichert wird ergeben sich für eine komplette kinetische und thermodynamische Analyse etwa 50-100 .tif Dateien. Da das Signal/Hintergrund Verhältnis bei den oberflächengebundenen Verfahren und das Rauschen durch die Quantisierung schwer zu erfassen ist sind Tiefpass Filterverfahren und ableitungsbasierte Schmelzpunktermittlung im Gegensatz zu lösungsbasierten Verfahren nicht zielführend. Tiefpassfilterung erhöht auch die Anzahl der notwendigen Messungen.

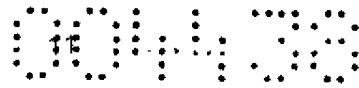
Die Verifikation des wurde mit 40 Experimenten gemacht, die Realisierung des Algorithmus wurde mit Matlab gemacht.

Für die Verifikation der Methode wurde der „Actin X-chip“ entwickelt (Krainer et al., 2010). Aktin ist eines der weitverbreitesten Gene in Eukaryoten mit Potential zur Lebensmitteldiagnostik (Goodson and Hawse, 2002; Hightower and Meagher, 1986). Durch den Vergleich von Gensequenz aus unterschiedlichen Spezien konnte ein breites Spektrum an Homologie realisiert werden. Die Fängermoleküle wurden in unterschiedlichen Längen realisiert. Als Proben wurden beispielhaft DNA und RNA verwendet.

Für das Probendesign wurde ARB verwendet (Ludwig et al 2004). Die Fängermoleküle wurden mit Aminogruppen modifiziert.

Die Fängermoleküle wurden in 10 µl 50% DMSO gelöst und in drei Replikaten gespottet. Für die 6 wichtigsten Proben wurden vier Konzentrationen gespottet (5.6, 16.7, 50.0 and 150 µM). Damit wurde der Einfluß der Dichte der Fängermoleküle auf die Kinetik und Denaturierung ermittelt. Es wurden slides der Firma Genewave verwendet. Gespottet wurde mit SMP3 stealth Nadeln auf einem Omnigrid 100 Spotter. Das postprocessing der Slides wurde wie in (Bodrossy et al., 2003) durchgeführt.

Es wurde die genomische DNA (gDNA) von folgenden Pflanzenarten extrahiert und amplifiziert *Kartoffel (Solanum tuberosum L.)* , *Paprika (Capsicum annuum L.)*, *Bohne*



(*Vicia faba*), Tabak (*Nicotiana glauca*), Heidelbeere (*Rubus idaeus*), Mais (*Zea mays* L.), Soja (*Glycine max*), Sonnentau (*Drosera rotundifolia* L.), Kastanie (*Castanea sativa* Mill.), Kalanchoe (*Kalanchoe daigremontiana*), Apfel (*Malus domestica*), Kiefer (*Pinus sylvestris* L.), Weizen (*Triticum aestivum* L.), Götterbaum (*Ailanthus altissima*), Raps (*Brassica napus*).

Um sicherzustellen dass nur ein einziges Genfragment amplifiziert wurde (McKinney and Meagher, 1998; Muller et al., 2007) wurden die PCR-Produkte mit einem TA cloning kit (Clontech) geklont. Die Plasmid DNA wurde von *E.coli* mittels alkalischer Lysierung extrahiert. Nach der Extraction und Amplifikation mittels PCR wurden die Fragmente sequenziert (AGOWA GmbH) und sind in Genbank submitted (Krainer S., 2009). Die Preparation der Proben ist in (Bodrossy et al., 2003) beschrieben.

Zur Kontrolle der PCR und zur Ermittlung der Fluorophore Inkorporation wurde ein ND-100 Spectrophotometer (NanoDrop, Wilmington, DE) verwendet.

Beispielsweise wurde folgender Versuch durchgeführt:

Die Hybridisierung erfolgte in einer Genewave Hybridisierungs- und Scanningstation (Marcy et al., 2008). Der Hybridisierungslösung bestand aus 6xSSC, 1x Denhardt's, 0.1 % SDS and 45% Formamid für die Systeme mit langen oligos und 25% formamide für die kurzen oligos. Die Konzentration der Proben wurde von 0.5 to 3.6 nM variiert. Die Lösung wurde für 5 min auf 95° erhitzt und danach sofort auf Eis gestellt um den Einfluss von Sekundärstrukturen zu reduzieren.

Während der 12 h Hybridisierungszeit wurde die Temperatur auf 42°C eingestellt und die Lösung agitiert. In der ersten Stunde wurde 1 Bild/Minute aufgezeichnet, danach alle 30 min ein Bild.

Nach der Hybridisierung wurden die slides bei 42°C gewaschen (5 min in 6x SSC, danach 5 min in 0.2x SSC). Vor dem Schmelzen wurden 15 Aufnahmen gemacht um sicherzustellen dass das System im Gleichgewicht ist und um das Ausbleichen des Fluorophors festzustellen. Zum Schmelzen wurde die Temperatur von 42°C to 85°C mit einer Rate von 1 °C/min erhöht. Dabei wurde alle 20s ein Bild aufgenommen.

Die Segmentierung und Quantifizierung der .tif files erfolgte mit der Hyblive software. Die Einstellung "irregular shape" wurde verwendet, alle anderen Parameter waren auf den Standardeinstellungen. Es wurde der Mittelwert aus den drei Replikaten genommen.

Patentansprüche:

1. Verfahren zur Approximation einer Schmelzkurve von in einer Probe befindlichen Moleküle, insbesondere Biomolekülen, wie z.B. DNA, RNA-, PNA-, LNA-Molekülen oder Proteinen, die an Fängermolekülen anhaften,

basierend auf einer Vielzahl von Messdatensätzen umfassend jeweils

- einen Messwert I, der proportional zur Konzentration der an einem Fängermolekül anhaftenden zu messenden Proteine ist, sowie
- einen zugehörigen gemessenen Temperaturwert T,

**dadurch gekennzeichnet, dass**

die vorgegebenen Messwerte mittels eines Fitting-Verfahrens an eine den Verlauf des Messwerts I für eine vorgegebene Temperatur T bezeichnende Kurve der Form

$$I(T) = \frac{k_a (k_{cy} - T)}{e^{k_h(T - k_{tm})} + 1} - k_{HT} * T + k_{const}$$

angepasst werden, und wobei die Fitting-Parameter

$k_a$ , und  $k_b$ ,  $k_{cy}$ ,  $k_{HT}$ ,  $k_{Tm}$  und  $k_{const}$  im Zuge des Fitting-Verfahrens an die aufgenommenen Messdatensätze angepasst werden.

2. Verfahren gemäß Anspruch 1, dadurch gekennzeichnet, dass sämtliche Messwerte einer linearen Transformation unterzogen werden, die diese auf Werte innerhalb des Intervalls [0, 1] abgebildet, wobei insbesondere der Wert I(Tmin) der Anfangstemperatur auf 1 und der Wert I(Tmax) auf null abgebildet wird.

3. Verfahren gemäß Anspruch 1 oder 2, dadurch gekennzeichnet, dass

- die Temperatur der Probe kontinuierlich erhöht wird, und
- die Probe gegebenenfalls bestrahlt und/oder mit Markermolekülen versetzt wird, und
- die Messdatensätze aufgezeichnet werden, indem in vorgegebenen Zeitabständen Messungen der Messwerte gemeinsam mit der zum selben Zeitpunkt gemessenen Temperatur (T) vorgenommen werden.

4. Verfahren gemäß Anspruch 3, dadurch gekennzeichnet, dass laufend jeweils nach der Bestimmung einer Anzahl von Messdatensätzen jeweils ein Fitting-Schritt durchgeführt wird und basierend auf den bisher ermittelten Ergebnissen jeweils eine verbesserte Fitting-Kurve erstellt wird.

5. Verfahren gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass die Probe in einen Probenträger verbracht wird, wobei an einer der Begrenzungsflächen des Probenträgers Fängermoleküle angeordnet sind, die eine

vorgegebene Art von Molekülen binden, wobei die Anzahl der an den Fängermolekülen befindlichen Moleküle, z.B. von DNA, RNA, PNA, LNA oder Proteine als Messwert I bestimmt werden.

6. Verfahren gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass als Messwert der von die Probe reflektierte und/oder, insbesondere durch Lumineszenz, vorzugsweise Fluoreszenz, abgestrahlte Lichtintensität herangezogen wird.

7. Verfahren gemäß Anspruch 6, dadurch gekennzeichnet, dass der Probenträger an zumindest einer Seite von einem Microarray begrenzt ist, und die Bestimmung des Messwerts durch Aufnahme eines Digitalbildes des Microarrays durchgeführt wird.

8. Verfahren gemäß Anspruch 7, dadurch gekennzeichnet, dass jedem Rasterpunkt des Microarrays vorab ein oder mehrere Pixel des Digitalbilds zugeordnet werden und aus den Lichtintensitätswerten der jeweiligen Pixel die jeweiligen Messwerte I der in den Rasterpunkten des Microarrays befindlichen Proben bestimmt werden und dass für jeden Rasterpunkt des Microarrays laufend separate Messungen und Fitting-Verfahrensschritte durchgeführt werden.

9. Verfahren gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass als Messwert die elektrische Leitfähigkeit, mechanische Eigenschaften, chemische Veränderungen, die Magnetisierbarkeit, die Impedanz bei einer vorgegebenen Frequenz, insbesondere zwischen 1 Hz und  $10^7$  Hz, oder die Änderung des Brechungsindex herangezogen wird.

10. Verfahren gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass als Messwert I die Gewichtszunahme bedingt durch das Anhaften der Moleküle an den Fängermolekülen herangezogen wird.

11. Verfahren gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass die Temperatur T mit einer Rate von 0,1 bis 5°C/min, insbesondere 0,8 bis 1,5°C/min erhöht wird.

12. Verfahren gemäß einem der vorangehenden Ansprüche, dadurch gekennzeichnet, dass zur Bestimmung der Fitting-Parameter  $k_a$ , und  $k_b$ ,  $k_{Cy}$ ,  $k_{HT}$ ,  $k_{Tm}$  und  $k_{const}$  ein Gauß-Newton-Verfahren verwendet wird.

13. Datenträger auf dem ein Programm zur Durchführung eines Verfahrens gemäß einem der Ansprüche 1 bis 12 abgespeichert ist.

14. Datenträger mit elektronisch auslesbaren Steuersignalen, die so mit einem programmierbaren Computersystem zusammenwirken können, dass ein Verfahren nach einem der Ansprüche 1 bis 12 ausgeführt wird.

15. Computerprogramm mit Programmcode-Mitteln eingerichtet zur Durchführung eines Verfahrens nach einem der Ansprüche 1 bis 12, wenn das Programm auf einem Computer ausgeführt wird,

16. Computerprogramm nach Anspruch 15, gespeichert auf einem Datenträger.

17. Computerprogrammprodukt mit Programmcode zur Ausführung des Verfahrens nach einem der Ansprüche 1 bis 12, wenn das Programm auf einem Computer ausgeführt wird.

Wien, am 28. April 2010



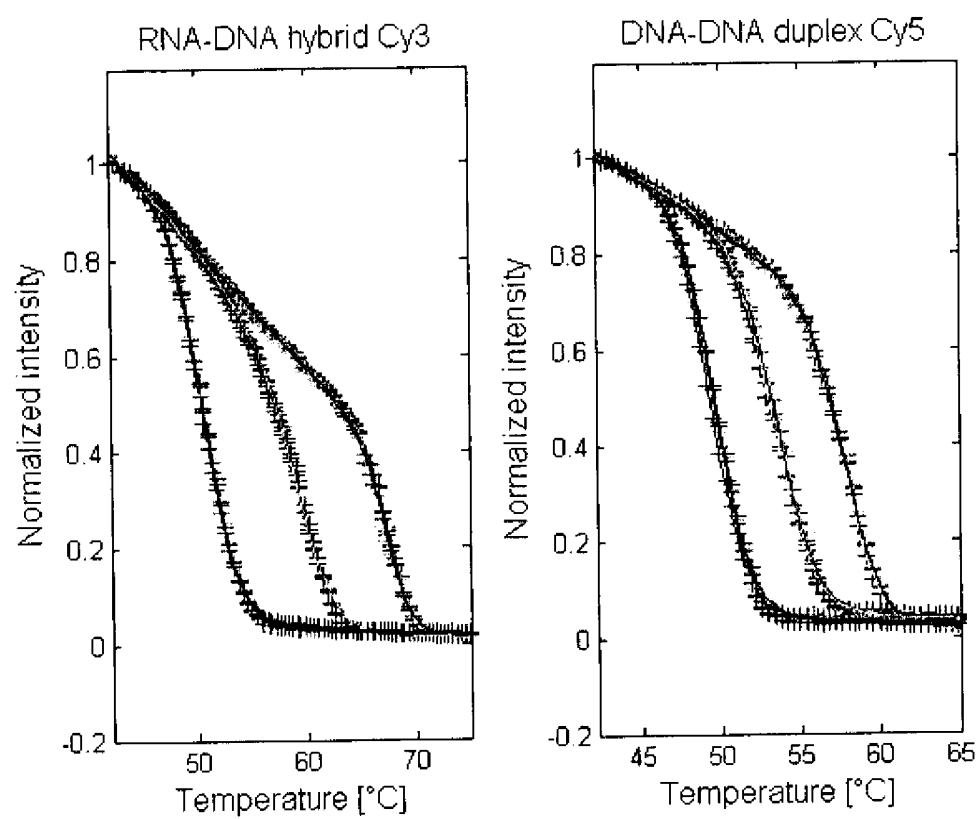


Fig. 1

004:38

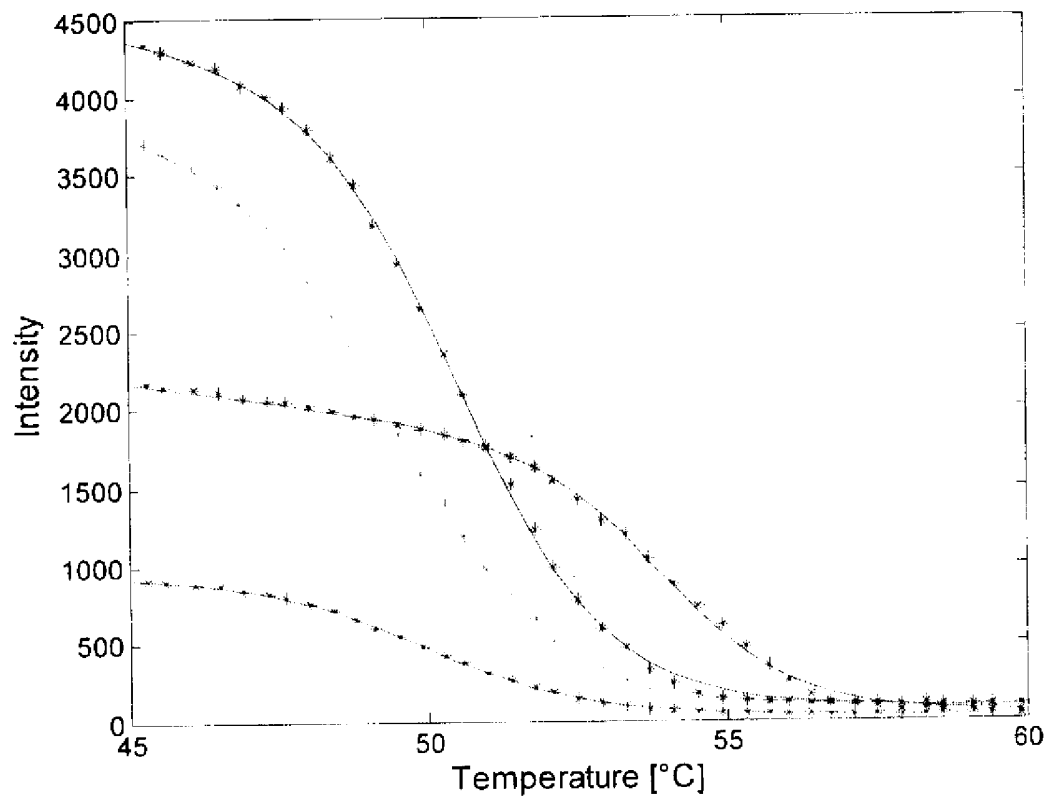


Fig. 2

004430

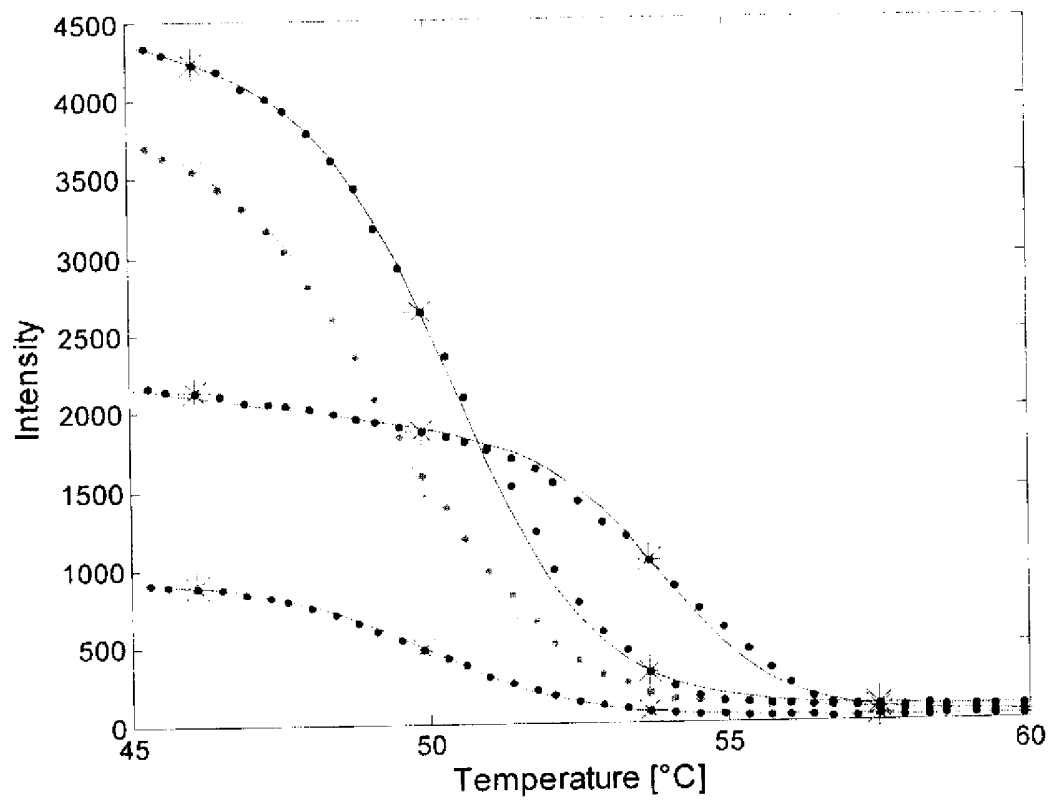


Fig. 3

001433

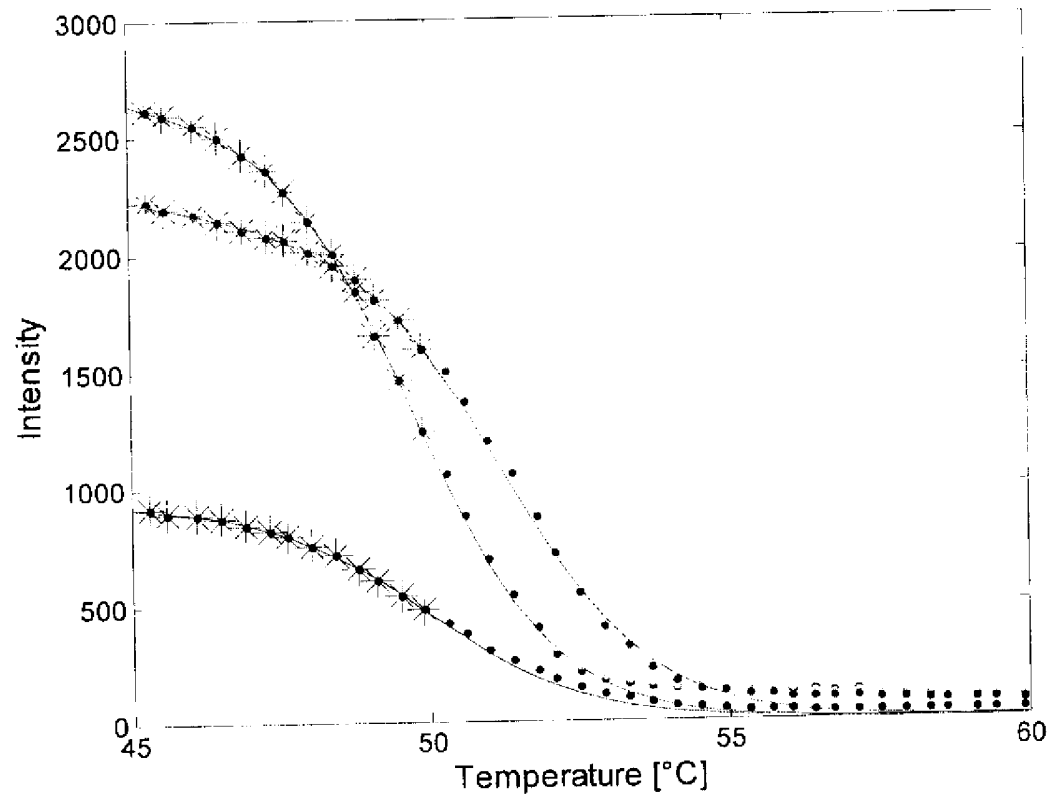


Fig. 4

0014:30

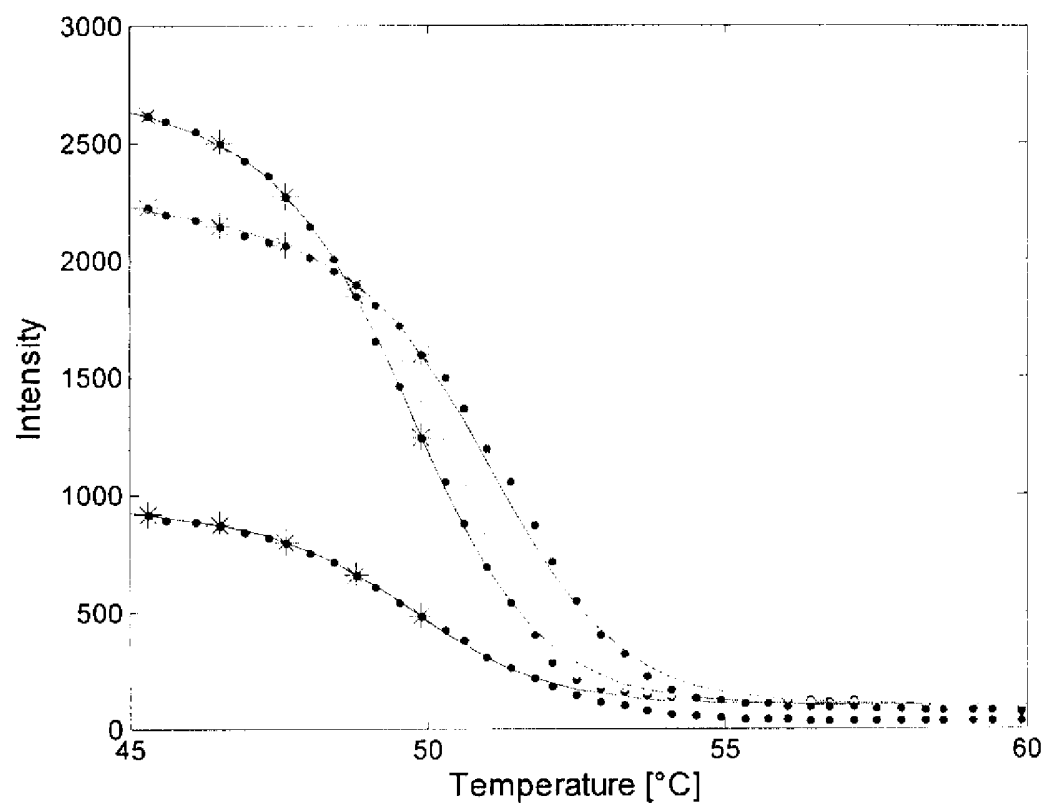


Fig. 5